# 进展情况 —— 技术部分

## 数据获取技术进展

目前主要以高校官网公布的讲座信息为主要信息来源，使用Python进行公开讲座信息的收集工作，辅以人工手段对文本数据进行进一步处理（包括提取关键信息）。收集的数据目前已上传至[Github仓库](https://github.com/JJYDXFS/little-innovation/tree/master/Text_Materials)。

## 网站搭建技术进展

项目最后的呈现方式为：微信公众号定期提供讲座分享+Web端网站提供讲座资源聚合服务及自然语言相关的应用，本部分主要侧重于说明WebApp的搭建进展。网站首页地址为<http://111.229.84.244:5000/>。目前使用html/css/js+Flask的前后端不分离形式进行网站demo的搭建，数据库使用MySQL8.0。前后端分工为由前端编写网页模板、后端Flask主要负责对模板进行渲染和对，动态路由也由Flask实现。

### 1. 服务器环境搭建

服务器使用腾讯云服务器标准型S2，配置为1核 2GB 1Mbps，操作系统Ubuntu Server 18.04.1 LTS 64位，硬盘大小50GB。进行过基础LAMP（Ubuntu18+Apache2+MySQL+Php7）环境搭建，前期曾使用该方式进行过前后端分离的开发。目前在网站默认端口80也同步了网站主页。

### 2. 前端

采用传统的html+css+js技术进行开发，基于[os-templates](https://www.os-templates.com/)的开源前端模板，根据项目需求对界面和功能进行改造。使用JQuery与服务器进行数据交换、实现动态加载；使用Vue.js简化DOM操作，后期考虑采用Vue组件化开发对网站前端结构进行升级。目前前端共编写了3个网页模板：首页模板、学科页模板、搜索结果页模板（其中首页模板编写进度较快，其余两个模板只有简单demo）。

#### 2.1 前端功能设计

1. 首页

* 首页预计实现如图所示功能，包括提供各类搜索功能和讲座预告信息，最终是否引入用户注册登录功能暂时未确定。
* 

1图 主页设计

1. 学科页

* 学科页预计将提供学科词云和相关的预告、笔记。学科页功能设计如下。
* 

2图 学科页设计

1. 搜索结果页

* 搜索结果页将根据关键词给出搜索的讲座信息，搜索结果将按匹配程度排序，并提供部分内容的预览。
* 

3图 搜索结果页设计

#### 2.2 模板实现进展

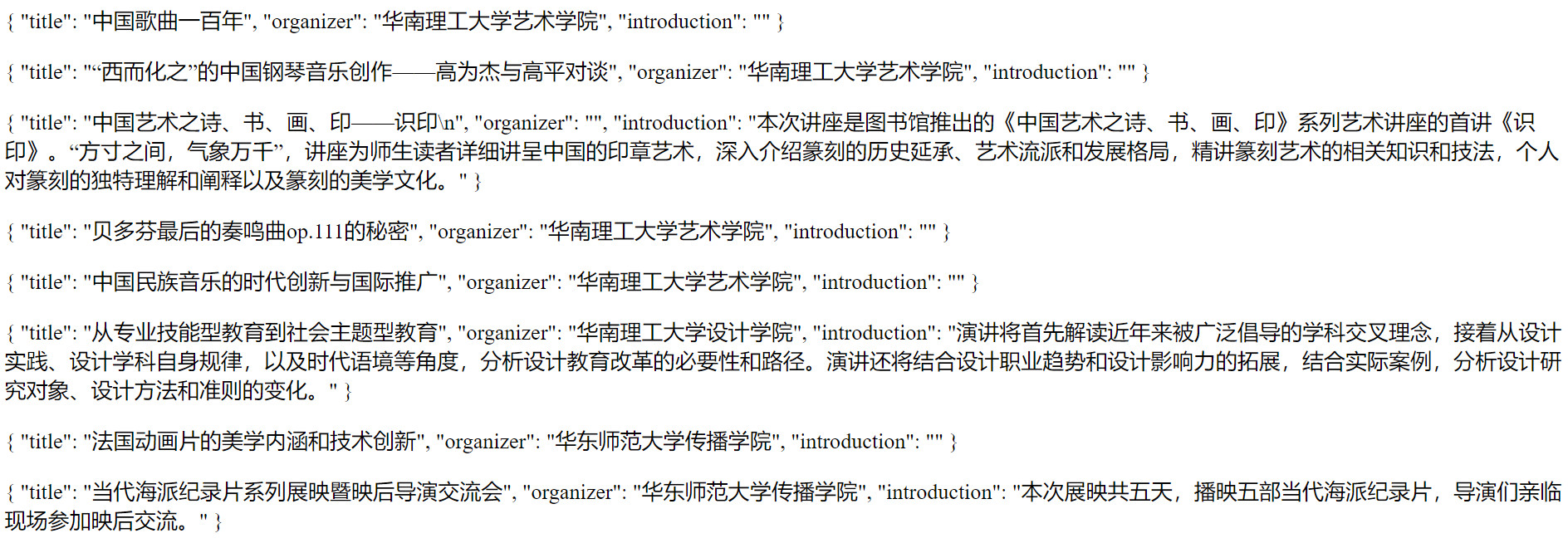
1. 网站首页模板

如下图所示，使用GET请求初步实现了导航栏的动态加载和点击跳转以及搜索框的自动跳转功能。



4主页截图

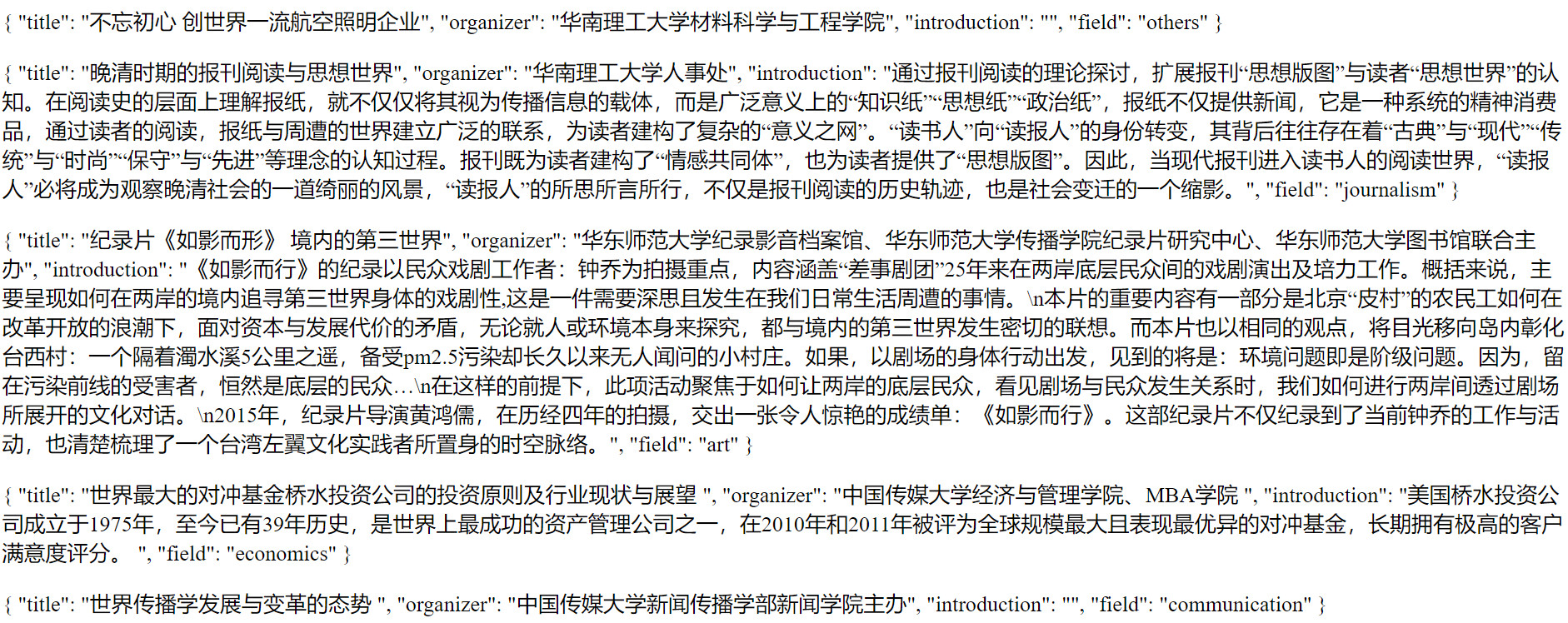
1. 学科页模板

* 学科页暂未进一步细化模板，直接使用POST请求向后端调取对应数据显示在页面上。如图示结果为“艺术”学科页的当前效果图。
* 

5“艺术”学科页截图

1. 搜索结果页模板

搜索结果页暂未进一步细化模板，直接使用POST请求向后端调取搜索结果（标题中含有关键词的讲座信息）显示在页面上。如图示结果为关键词“世界”的搜索结果。

* 

6关键词“世界”搜索结果

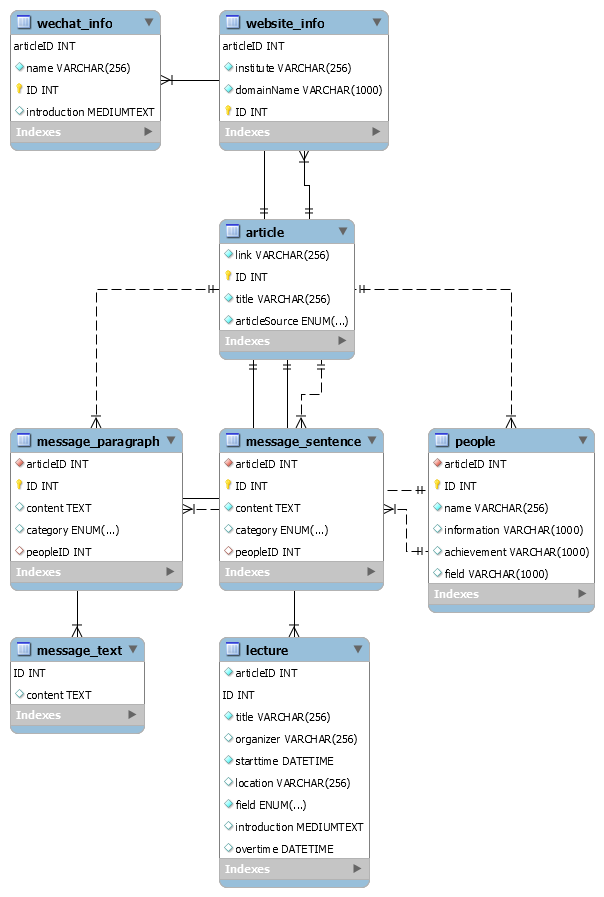
### 3. 后端

使用Python Flask框架，目前实现了学科页和搜索结果页的动态路由、上述三类模板页的渲染以及三个接口：

* 1. GET：获取学科列表
  2. POST：根据学科返回该学科包含的讲座信息
  3. POST：根据关键词返回标题中含关键词的讲座信息

### 4. 数据库

数据库ER图如图所示，目前数据库中一共收录讲座数据1048条，涵盖学科12类。



7 数据库ER图